

up the hierarchy

Claude Bennett

How productive are Extension programs? Government officials, legislators, university administrators, and Extension leaders, including you, want to know. Evidence of Extension's effectiveness can help decision makers improve service to clientele.

Chain of Events

Let me suggest seven categories of criteria for evaluating Extension programs and offer guidance in choosing evidence in these categories. These categories are based on a seven-link "chain of events."¹

First in the chain is *inputs*, the resources expended by Extension. These inputs produce *activities* that *involve people* who have *reactions*, pro and con. People involved may change their *knowledge, attitudes, skills, and aspirations (KASA)*. *Practice change* occurs when people apply their KASA change to working and living. What follows from these practice changes are *end results*. Such results should include accomplishing ultimate aims of the Extension program. Is there more profit because of adoption of new herbicides, or better health because people are beginning to eat a recommended diet?

These seven events are shown as a hierarchy in Figure 1. This "staircase" reaches toward solving, at the seventh level, one or more clientele problems.

The following are examples of criteria at each level of the hierarchy.

At the *inputs* level, criteria are plans (objectives) to allocate certain resources to a program, such as:

- time of paid staff and volunteers ("five staff-years will be allocated to the program").
- staff qualifications—paid and volunteer ("all program assistants must be neighborhood opinion leaders").

At the *activities* level, criteria are plans to perform, through the above inputs, certain educational activities, such as:

Claude Bennett: Specialist, Educational Methodology and Evaluation, Staff Development, Extension Service-USDA.

- publicizing programs (“five newspaper releases will be prepared”).
- transmitting subject matter through mass media, meetings, and other events (“five TV stations will show a video tape on this subject”).

At the *people involvement* level, criteria are plans for certain types and number of persons, groups, or communities to be involved:

- number of participants in programs, tours, meetings, or clubs (“200 4-H participants from low-income families will be enrolled in this project”).
- continuity, frequency, and intensity of interaction between clientele and Extension (“attendance at homemaker meetings will average 80% of membership”).

At the *reactions* level, criteria are plans to obtain certain reactions to involvement in activities, such as:²

- interest in educational events (“75% positive reactions to this topic”).
- acceptance of leaders (“leader of meeting rated competent by 2/3 of audience”).

Extension program evaluation isn't an end in itself. It's worth doing only if it helps in making decisions about program continuation, priorities, modifications, and so on.

At the *KASA change* level, criteria are plans that certain *knowledge, attitudes, skills, and aspirations* (KASA) will ensue from participation in activities:³

- direction (content) and extent of KASA change (“80% of homemakers will be able to select the most suitable furniture arrangement for homes”).
- durability of any KASA change (“95% of participating farmers, will still recall sources of pesticide safety rules 1 year after workshop”).
- intensity of attitudes to be acquired (“all youth in the citizenship seminar should disapprove nonvoting by the close of the seminar”).
- height of aspiration (“each couple in the workshop should decide to prepare a legal will within one month”).

At the *practice change* level, criteria are plans for changes in practices, technology, or social structure, as a consequence of KASA change, in terms of:

- individual innovation and adoption (“90% of farmers to adopt new, superior variety of wheat within 2 years”).

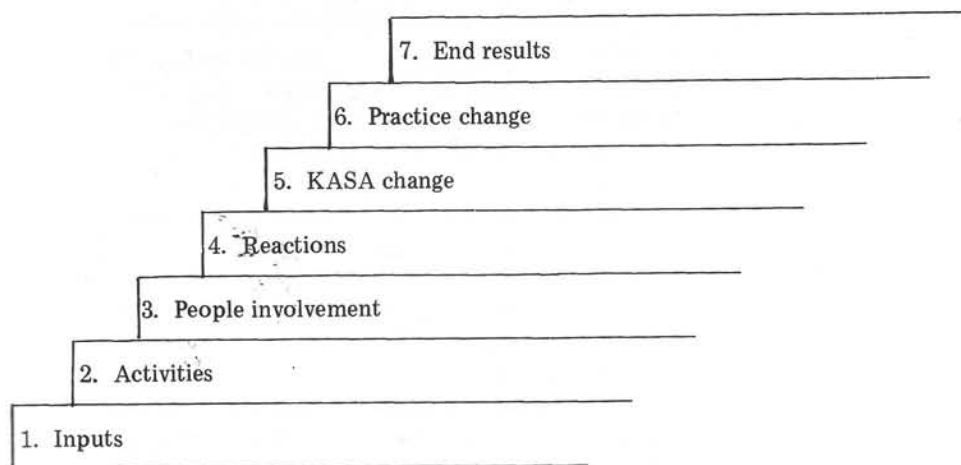


Figure 1. A hierarchy of evidence for program evaluation.

- collective change (“100% of communities to establish land-use planning boards within 4 years”).

At the *end results* level, criteria are plans that certain ultimate effects will come through change in practice:

- individuals (“one-third of ‘isolate’ youth attending camp to gain increased peer and self-acceptance”).
- groups (“the community will increase to 5% its annual rate of real economic growth”).

Unexpected “side” effects shouldn’t be ignored at any level of the hierarchy, but apply especially to end results. Side effects may be beneficial (serendipity) or harmful (“backlash”). For instance, new industry brought to a community with Extension’s help may alter established social relationships in unexpected ways.

Reaching Objectives

The more nearly program objectives or plans are reached, the more people value the program.⁴ A highly valued program is more likely to get continued or additional funding, as its objectives are likely to be retained, intensified, or broadened.

Selecting Level and Quality of Evidence

An Extension program usually has several objectives. These may be at several or all levels of the hierarchy shown in Figure 1. At which of the seven levels should you try to get evidence of program accomplishments? Three guidelines may help here.

Guideline 1: Evidence of program impact becomes stronger as the hierarchy is ascended. The two lowest levels provide little or no measure of clientele benefits. Level 3 provides one way of measuring opportunity for education to occur.

Ascending to the fourth level, reactions, provides somewhat better confirmation of whether activities are helpful. It's usually desirable to have at least evidence of the extent to which objectives for KASA change are reached. But KASA changes are only stepping stones to adoption of recommended practices.

Ideally, impact of more Extension programs would be evaluated in terms of whether desired end results are achieved.

But, here's the rub. *Guideline 2: The difficulty and cost of obtaining evidence of accomplishments increases as the hierarchy is ascended.* Resources required to measure actual outcomes generally increase as the hierarchy is climbed due to (a) increased scattering of sources of evidence, (b) greater time-lag following program activities, and (c) increased

Table 1. Examples of "hard" and "soft" evidence in a hierarchy for program evaluation.

Levels	Examples	
	"Hard" evidence	"Soft" evidence
7. End results	profit-loss statements; life expectancies and pollution indexes	casual perceptions of quality of health, economy, and environment
6. Practice change	direct observation of use of recommended farm practices over a series of years	retrospective reports by farmers of their use of recommended farm practices
5. KASA change	changes in scores on validated measures of knowledge, attitudes, skills, and aspirations	opinions on extent of change in participants' knowledge, attitudes, skills, and aspirations
4. Reactions	extent to which random sample of viewers can be distracted from watching a demonstration	recording the views of only those who volunteer to express feelings about demonstration
3. People involvement	use of social participation scales based on recorded observations of attendance, holding of leadership positions, etc.	casual observation of attendance and leadership by participants
2. Activities	pre-structured observation of activities and social processes through participant observation, use of video and audio tapes, etc.	staff recall of how activities were conducted and the extent to which they were completed
1. Inputs	special observation of staff time expenditures, as in time-and-motion study	staff's subjective reports about time allocation

probability of impact by sources of change other than the program.

Guideline 3: While hard evidence is usually ideal, it's more expensive and difficult to obtain.

Evidence is often referred to as "hard" versus "soft." Evidence is "hard" to the extent that it reflects precisely true characteristics of individuals, groups, or situations.⁵

Table 1 shows examples of hard and soft evidence at each level of the hierarchy. Hard and soft constitute a continuum, but the two categories are used for definition.

Guiding Decision Making

Extension program evaluation isn't an end in itself. It's worth doing only if it helps in making decisions about program continuation, priorities, modifications, and so on.⁶

When you select evidence for evaluating Extension programs, ask these questions:

1. Which levels of the hierarchy (Figure 1) contain the kinds of evidence we need to make decisions?
2. How "hard" does the evidence need to be?
3. Are resources available to obtain the level and hardness of evidence needed?
4. If resources are insufficient to provide the level and hardness of evidence desired by decision makers:
 - a. Will decision makers be able to use a lower level of evidence, or softer evidence?
If not:
 - b. Can you get more resources?

Without clear criteria, evidence of program accomplishments won't indicate the success of your Extension program. But, you can't judge program success or make sound decisions without enough evidence of program accomplishments.

Footnotes

1. Several elements of the chain have been identified by Kirkpatrick and Suchman. See Donald L. Kirkpatrick, "Evaluation of Training," in *Training and Development Handbook*, Robert L. Craig and Lester R. Bittel, eds. (New York: Mc-Graw-Hill, 1967), pp. 87-112 and Edward A. Suchman, *Evaluation Research* (New York: Russell Sage Foundation, 1967).
2. It doesn't matter initially why participants are interested. They may be attracted to discussions on grain production, etc., because they like to talk, not because of any particular interest in the subject; audiences may begin to watch an Extension television program just because there are pretty girls on it.
3. At levels 5, 6, and 7, whose objectives—Extension's or clientele—may become an issue. Degree of consensus on objectives at these levels will depend on adequacy of Extension program planning.
4. Comparing the extent to which objectives and achievements coincide isn't the only way to evaluate program impact. Other approaches

- include observing what was achieved in relation to program expenditures, and comparing achievements of similar programs. See Sara M. Steele, *Six Dimensions of Program Effectiveness* (Madison: University of Wisconsin-Extension, Division of Program and Staff Development, 1972). Also see Robert E. Stake, "The Countenance of Educational Evaluation," *Teachers College Record*, LXVIII (April, 1967), 523-40.
5. For an introduction to validity, quantification, and representativeness of evidence, see Claire Selltiz and others, *Research Methods in Social Relations* (New York: Holt, Rinehart and Winston, 1961).
 6. Daniel L. Stufflebeam, "Toward a Science of Education Evaluation," *Educational Technology*, VIII (July, 1968), 5-12; Joseph S. Wholey and others, *Federal Evaluation Policy: An Overview* (Washington, D.C.: Urban Institute, 1970); and W. Keith Warner, "Feedback in Administration," *Journal of Cooperative Extension*, V (Spring, 1967), 35-46.